IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of: TI-32328

Yifan Gong

Examiner: TBD

Serial No: TBD

Art Unit: TBD

Filed: Herewith

For: Method of Speech Recognition with Compensation for Both Channel Distortion and Background Noise

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D. C. 20231

"EXPRESS MAILING" Mailing Label No. EL547743702US I hereby certify that this paper is being deposited with the U.S. Postal Service Express Mail Post Office to Addressee Service under 37 CFR 1.10 on the date shown below and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Robert L Troike 1/18/02
Robert L. Troike, Reg. No. 24,183 Date

Dear Sir:

This application claims priority under 35 USC 119 (e)(1) of provisional application number 60/275,487, filed 03/14/01.

Prior to the examination of the above-identified application, please amend the specification by inserting before the first line the sentence:

--This application claims priority under 35 USC § 119(e)(1) of provisional application number 60/275,487, filed 03/14/01.—

IN THE SPECIFICATION:

Page 8, lines 17-19

A subset of WAVES database containing hands-free recordings was used, which consists of two recording sessions: parked (car parked, engine off) and city-driving (car driven on a stop and go basis).

REMARKS

The changes made on page 8 are illustrated in the attached "Version with markings to show changes made".

Respectfully submitted,

Robert L. Troike
Attorney for Applicant
Reg. No. 24,183

Texas Instruments Incorporated
P. O. Box 655474, MS 3999
Dallas, TX 75265
(202) 639-7710
Fax: (202) 639-7890

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE SPECIFICATION:

Page 8, lines 17-19

A subset of WAVES database containing hands-free recordings was used, which consists of [three] two recording sessions: [parked-trn (car parked, engine off),] parked (car parked, engine off)[,] and city-driving (car driven on a stop and go basis).

2. Use equal probabilities for $P_{\mathcal{H}}(p)$, $P_{\mathcal{J}\,|\,,}P(j\,|p)$ and $P_{\mathcal{K}|\,,,\,\mathcal{J}}(k\,|\,h,j)$.

$$
\begin{aligned}
P_{\mathcal{H}}(p) &= \mathrm{C} \\
P_{\mathcal{J}|H}(j\,|p) &= \mathrm{D} \qquad (10)\\
P_{\mathcal{K}|H,,\mathcal{J}}(k|p,j) &= \mathrm{E}
\end{aligned}
$$

3. In fact, the case described in Eq-10 consists in averaging the compensated mean vectors $\overline{\mathrm{m}}_{p,j,k}$. Referring to Eq-4 and Eq-1, it can be expected that the averaging reduces the speech part $\mathrm{m}_{p,j,k}$ just as CMN does. Therefore, Eq-7 could be further simplified into:

$$
\hat{\mathrm{b}} = IDFT\,(DFT(\mathbf{b}) \oplus DFT\,(\widetilde{\mathrm{X}})). \qquad (11)
$$

The model $\mathrm{m}^{T}_{p,j,k,t}$ is then used with CMN on noisy speech. Unfortunately, $\hat{\mathrm{b}}$ is a function of both channel and background noise in all above cases. In other words, in presence of noise, there is no guarantee that the channel will be removed by such a vector, as is for CMN.

A subset of WAVES database containing hands-free recordings was used, which consists of two recording sessions: parked (car parked, engine off) and city-driving (car driven on a stop and go basis).

In each session, 20 speakers (10 male) read 40 sentences each, giving 800 utterances. Each sentence is either 10, 7 or 4 digit sequence, with equal probabilities. The database is sampled at 8kHz, with MFCC analysis frame rate of 20ms. Feature vector consists of 10 statis and 10 dynamic coefficients.

HMMs used in all experiments are trained in TIDIGITS clean speech data. Utterance-based cepstral mean normalization is used. The HMMs contain 1957 mean vectors, and 270 diagonal variances. Evaluated on TIGIDIT test set, the recognizer gives 0.36% word error rate.